DOCUMENT RESUME

ED 044 546 AA 000 648

TITLE	Teacher Evaluation. PREP-21.
INSTITUTION	National Center for Educational Communication (DHEW/OE), Washington, D.C.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	PREP-21
PUB DATE	Jan 71
NOTE	57p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$2.95
DESCRIPTORS	*Decision Making, Effective Teaching, *Evaluation Techniques, Measurement Instruments, *Personnel Evaluation, *Teacher Evaluation
IDENTIFIERS	*Putting Research into Educational Practice (PREP)

ABSTRACT

Dr. Dale L. Bolton of the University of Washington, Seattle, in an interpretive studies project for the Office of Education, surveyed evaluation practices in business and industry, Government agencies, and some of the larger school systems in the country. In his final report on this project, "Selection and Evaluation of Teachers," Dr. Bolton elaborates on the above findings and makes implications for the educator. The topic of "Recruitment and Selection of Teachers" was covered in PREP kit No. 20. PREP kit No. 21, "Teacher Evaluation," is also based on his final report. Both kits and the final report are being entered into the ERIC system. Copies of the kits will be available in microfiche and hard copy from the ERIC Document Reproduction Service (EDRS). The final report will also be available from EDRS, but only in microfiche. For related document see ED 043 797. (Author/LS)

AA 000 648

ED 044 546

PROBLEM ► RESEARCH ►

INTERPRETIVE REPORTS OF EDUCATIONAL RESEARCH AND DEVELOPMENT

Enclosed are specially designed materials on a topic of current interest to educators. The purpose of the materials, produced under U.S. Office of Education contracts, is to bring research and development findings to bear on the practical problems of educators.

Because OE is able to produce only a limited number of copies, the materials are designed so that educators can easily and inexpensively reproduce or adapt them to meet local needs, and distribute them in their educational communities. Other studies are being supported on problems now facing school personnel. As materials from these studies become available, they will be disseminated in the same manner.

PREP BRIEF

No. 21

TEACHER EVALUATION

Teacher evaluation can run the gamut from opinion to precise measurement. Administrators must first decide their purpose(s) for evaluating teachers, and then determine how precise they want to be in their evaluations. Teachers may be evaluated:

- To improve teaching
- To reward superior performance
- To supply information for modifying assignments
- To protect either the individual or the organization in legal matters
- To validate the selection process
- To provide the basis for planning for individual growth and development

Most educators consider the improvement of instruction to be the most important purpose for evaluation. However, the others listed above should also be considered in designing evaluation procedures. Procedures are most effective when they are cooperatively planned by teachers and administrators who receive assistance from specialists, consultants, parents, and students. The most effective evaluation plan is one which improves learning conditions for students and contributes to higher teacher morale.

Steps in the evaluation process include:

- Goal setting conferences
- Observation and information collection
- Post-observation conferences, communication
- Decisionmaking
- Assessment of the evaluation process

Selecting an adequate and competent measuring instrument is crucial to the evaluation process, and the following should be considered:

- Its relevance and validity--Does it measure what it is intended to measure?
- Its reliability--Does it continue to maintain its stability from one application to another?
- Its fidelity--Does the response to the instrument parallel the actual performance?
- Its ease of administration and scoring--How much time is needed to administer the instrument? Is it easily scored or interpreted?
- Its cost--Is it practical and worth the cost?
- Its "taboo" factor--Does it conflict with local customs or traditions?

Problems are encountered in teacher evaluation when there is an emphasis on fault-finding rather than on helping teachers; when prejudice, bias, and poor judgment are used in collecting and analyzing information; and when communication is not open and honest.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

FOR MORE INFORMATION

Dr. Dale L. Bolton of the University of Washington, Seattle, in an interpretive studies project for the Office of Education, surveyed evaluation practices in business and industry, Government agencies, and some of the larger school systems in the country. In his final report on this project, "Selection and Evaluation of Teachers," Dr. Bolton elaborates on the above findings and makes implications for the educator. The topic of "Recruitment and Selection of Teachers" was covered in PREP kit No. 20. PREP kit No. 21, "Teacher Evaluation," was also based on his final report. Both kits and the final report are being entered into the ERIC system. Copies of the kits will be available in microfiche and hard copy from the ERIC Document Reproduction Service (EDRS), 4936 Fairmont Boulevard, Bethesda, Maryland 20014. The final report will also be available from EDRS, but only in microfiche.

PREP

No. 21

PREP is . . .

- *a synthesis and interpretation of research, development, and current practice on a specific educational topic*
- *a method of getting significant R&D findings to the practitioner quickly*
- *the best thinking of researchers interpreted by specialists in simple language*
- *the focus of research on current educational problems*
- *a format which can be easily and inexpensively reproduced for wide distribution*
- *raw material in the public domain which can be adapted to meet local needs*
- *an attempt to improve our Nation's schools through research*
- P *utting*
  R *esearch into*
  E *ducational*
  P *ractice*

TEACHER EVALUATION

All teachers *are* evaluated. Regardless of how formal the system for evaluation is, what evidence is collected or analyzed, how often formal reports are written--teachers are evaluated and they are evaluated rather often. Students, parents, other teachers, administrators and supervisors, and even the public evaluate teachers. The question facing both administrators and teachers, then, is not whether teachers *should* be evaluated, since this cannot be avoided; rather the question is how systematic the evaluation should be in order to be most effective.

Effectiveness must be in terms of certain purposes desired for the school district; and the design of the evaluation system should include ways of collecting and processing information, communicating with the people concerned, making decisions, and assessing how well the evaluation system works.

This kit is the second one adapted from the final report of the interpretive studies project "Selection and Evaluation of Teachers" conducted by Dr. Dale L. Bolton of the University of Washington, Seattle. Some of the questions answered in the 11 documents comprising the kit are:

What are the purposes of teacher evaluation?
Who should be involved in planning for evaluation and in the evaluation itself?
How should evaluators be trained?
What decisions should be made on the basis of evaluation?
How does one select an instrument for measuring teacher behavior?
How can the evaluation process be analyzed for improvement?

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

This kit is being entered into the ERIC system, and copies will be available in microfiche and hard copy from the ERIC Document Reproduction Service (EDRS), 4936 Fairmont Boulevard, Bethesda, Maryland 20014. The final report upon which this kit was based will also be in the system, but will be available from EDRS in microfiche only.

IPRIP

No. 21-A

EVALUATION--AN OVERVIEW

Many people consider the evaluation of teachers to be mostly a matter of opinion. The problem underlying this viewpoint is the question of precision--how precise a principal desires evaluation to be, how precise the superintendent and school board desire evaluation to be, and how precise teachers desire evaluation to be. What are some of the implications of varying degrees of precision?

For example, suppose you wanted to measure the length or the circumference of a small object. Three different tools might be used: a piece of string, a ruler, or a micrometer.

The string has obvious advantages and disadvantages: it is cheap, plentiful, easily available, expendable, and easy to use or explain to an untrained person. On the other hand, it tends to wear out and stretch with age and different people tend to stretch it to differing lengths when they use it for measuring.

The ruler has definite advantages over the string: it does not stretch (increasing what is called the reliability of the measurement) and it tends to last longer than the string. However, rulers are not so available as string, and untrained people make mistakes with them.

The micrometer is much more expensive than either of the other tools, is less accessible, and is difficult for some people to use. Yet, it has an accuracy that is necessary for some jobs--an accuracy that cannot be obtained with the other tools--and in the hands of a trained operator it is seldom misused.

The analogy to measurement in the evaluation process appears clear. Some measurement devices have characteristics similar to the string; others, while more accurate, have some of the disadvantages of the ruler; while still others have the specialized uses of the micrometer. Therefore, measurement devices and procedures are chosen according to how precise one desires to be, and what types of resources are available; considerable precision can be obtained if it is desired.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

Measurement vs. Assessment

Note the distinction between measurement and evaluation. *Measurement* is many times used synonymously with *assessment* and is related to the qualified or quasi-quantified "description" of events, behavior, or outcomes. However, evaluation (in the context of evaluating teachers) has to do with "judgments" relating to the "goodness" of teacher behavior and/or results of that behavior in light of agreed-upon objectives. These objectives usually are approved (either tacitly or explicitly) by a particular school system and community, giving the objectives a type of social validity for that community.

Evaluate Program or Individuals?

There is a difference between assessing or evaluating the *program* of a school or a school system and evaluating the *individuals* who contribute to that system. The former attempts to look at the output of the total system, the procedures used to accomplish those outcomes, and the way those procedures were implemented. The latter examines the contribution to the system in terms of individual output, procedures, and implementation. This report is concerned with the evaluation of *teachers* rather than evaluation of *systems of teaching* within school systems.

The advantage to examining the effectiveness of teachers is that one can make better inferences about system performance from combinations of subsystem performance than he can make about subsystem performance from the total system performance. For example, if one knows something about each individual teacher's performance in a building, he can infer something about the total performance of the school. However, knowledge of the total productivity of the school would provide little basis for inferring how a given teacher behaves.

What is Involved

Accurate *measurement* can occur without *evaluation;* however, it is impossible for evaluation to be good without adequate measurement. In addition, the data collected via the measurement process must be adequately analyzed and interpreted if good evaluation is to occur. Therefore, *good evaluation is preceded by:*

- A determination of what is important (criteria)
- Measurement
- Analysis
- Interpretation

Judgments made prior to these activities are likely to be unsound. (Note that it is assumed that the making of judgments regarding teachers is inevitable.) The real issues are concerned with whether the *criteria are appropriate* and the *data are sound.*

⇨ CAUTION

In evaluating teachers, the emphasis is on making judgments in relation to objectives, not on judging the personal worth of people.

(viewpoint: research staff)

Although the research by Rose provides some indication of the nature of the resistance by teachers, what is the basis for the resistance by persons doing the evaluation? Interviews with persons involved indicate a number of factors may be present. For example:

- A general lack of certainty regarding criteria, measurement process and procedures for analysis and interpretation of data.
- A resistance to placing oneself in the position of manipulating or adversely affecting other people's lives.
- A fear of precipitating an unpleasant reaction on the part of the person being evaluated. The reaction is then said to prevent a relationship that is conducive to helping the individual improve.
- A lack of ability to cope with the weaknesses of the individual in terms of organizational needs and his ability to improve. This is sometimes linked with a failure to communicate to the individual the necessity of dealing with both the individual's and the organization's problems.
- A failure to see the relationship of evaluation of others to the purposes of the person doing the evaluation.
- An inability to organize time in such a manner that adequate observations can be made.

Models of Evaluation

What should be the model on which evaluation is based? Should it be modeled after a marriage contract ("for better, for worse," i.e., a tenure system) or after a professional baseball contract (no results, no renewal)? Would evaluation decisions and processes be any different if the teacher and principal had to agree on a contract annually?

⇨ SUGGESTION

Models of evaluation function within the constraints of conditions of employment; more flexible models may develop if these constraints are modified. (practice: industry)

Process and Product

Some systems examine only the outcomes of teaching, ignoring what processes have been used. Other systems emphasize the classroom procedure, ignoring what has been learned by pupils. A self-correcting system should be designed which will identify errors and make changes in procedure before harmful effects occur.

Formal evaluations should be analytic rather than comparative, establishing whether the teacher reaches various standards but avoiding attempts to compare the teacher with other teachers. The emphasis should be on helping individuals improve their contributions to the learning of school children rather than on taking punitive or controlling actions, making odious comparisons, or using questionable motivation techniques.

Reactions to Evaluation

There is some evidence (Rose, 1963) [1] that teachers welcome evaluation if:

- The major focus is on improving rather than fault finding
- The information produced is meaningful to the teacher
- The principal takes the necessary time to collect information that is adequate and to discuss it with the teacher

This research implies the need to have agreement on purpose and procedures, and this can occur only when purposes and procedures are specific. Also communications must be honest as well as skillful, and disagreements must be handled in the open rather than avoided.

⇨ NOTE

Changes in teacher or supervisor might be precipitated as a result of evaluation, and both the evaluator and the teacher are vulnerable when changes are possible. (position: research staff)

Considerable resistance to discriminatory evaluation of teachers exists in spite of the advantage of such evaluation.

⇨ CONCLUSION

Removal of resistance to evaluation depends on clear organizational goals, resources adequate for training evaluators (and providing adequate time for them to perform the tasks required), and clarity of the relationship of the organizational goals and the task of the evaluator. (practice: industry, schools)

Changes in goals and procedures of evaluation are resisted by various forces, and teachers' organizations are one of the strong forces opposed to discriminating evaluations or evaluations which expand beyond the single purpose of improving instruction.

---

[1] See document No. 21-J, "Bibliography," for references cited in the text

Although the research by Rose provides some indication of the nature of the resistance by teachers, what is the basis for the resistance by persons doing the evaluation? Interviews with persons involved indicate a number of factors may be present. For example:

- A general lack of certainty regarding criteria, measurement process and procedures for analysis and interpretation of data.
- A resistance to placing oneself in the position of manipulating or adversely affecting other people's lives.
- A fear of precipitating an unpleasant reaction on the part of the person being evaluated. The reaction is then said to prevent a relationship that is conducive to helping the individual improve.
- A lack of ability to cope with the weaknesses of the individual in terms of organizational needs and his ability to improve. This is sometimes linked with a failure to communicate to the individual the necessity of dealing with both the individual's and the organization's problems.
- A failure to see the relationship of evaluation of others to the purposes of the person doing the evaluation.
- An inability to organize time in such a manner that adequate observations can be made.

Models of Evaluation

What should be the model on which evaluation is based? Should it be modeled after a marriage contract ("for better, for worse," i.e., a tenure system) or after a professional baseball contract (no results, no renewal)? Would evaluation decisions and processes be any different if the teacher and principal had to agree on a contract annually?

⇨ SUGGESTION

Models of evaluation function within the constraints of conditions of employment; more flexible models may develop if these constraints are modified. (practice: industry)

Process and Product

Some systems examine only the outcomes of teaching, ignoring what processes have been used. Other systems emphasize the classroom procedure, ignoring what has been learned by pupils. A self-correcting system should be designed which will identify errors and make changes in procedure before harmful effects occur.

⇨ NOTE

The purpose of examining outcomes of teaching is to determine whether goals have been met; the purpose of examining procedures is to determine whether a specified plan is being followed; the purpose of comparing outcomes and procedures is to determine whether the procedures should be modified. (practice: industry, schools)

Goal Setting

An interview with the employee at the end of the first 6 months of employment may be used to set goals for the following year. This type of activity helps to coordinate the functions of selection and evaluation by integrating the personal plan for development with the evaluation procedure and also helps to express the organization's philosophy. Then the evaluator can help establish a training program for the individual, rather than simply correct faulty behavior. In this way the evaluation is used in a *positive* manner to plan and direct the individual's growth.

⇨ NOTE

Coordination of the personal and professional goals of a teacher leads to improved morale and productivity. (practice: schools)

Both the educational organization and individuals within it are being asked by the general public to be "accountable" for their actions and their products.

⇨ NOTE

Accountability consists of providing evidence regarding the degree of accomplishment of prespecified goals and objectives. Too often, if a child does not learn, the child, society, or some factor other than the teacher is blamed. Teachers must develop a sense of responsibility for providing this evidence as well as accomplishing goals. (position: research staff)

Satisfaction with "accountability" probably is related to:

- The degree of specificity of the goals
- Agreement on the goals
- The *perceived* adequacy of the evidence provided

A direct result of a desire for accountability is an increase in the number of States requiring annual evaluation of all (not just probationary) teachers.

PREP

No. 21-B

PLANNING FOR TEACHER EVALUATION

Suppose you had the responsibility for planning a teacher evaluation program for a school or a school district. Where would you begin? Whom would you involve? What decisions would need to be made? How rapidly would you try to implement the procedures?

Objectives

One of the first steps in planning a teacher evaluation program is to determine what is considered important in teaching--in the situation for which the evaluation program is being considered. This determination allows one to:

- Establish what is expected of teachers (as far as behavior and outcomes of behavior are concerned)
- Determine ways to measure what is happening
- Design ways to compare what occurs with what is desired

People Involved

Who will be involved in deciding what is important in teaching? Certainly, one should involve teachers in this decision and in the total design of evaluation procedures. The involvement of teachers in planning evaluation procedures is based on the premises that:

- A better plan will develop
- There will be more teacher commitment to the procedures
- Teachers will know what they are to do
- Teachers will know what will be evaluated

These premises are more than assumption or mere hopes. School districts which have involved teachers in planning have found that output and procedural goals for individual teachers are better understood and attained when they are cooperatively developed and written in precise terms than when they are unilaterally determined or written in very general language. Also, there is more commitment to goals and procedures when goals are specific and attainable, and when people who are to accomplish the goals are involved in establishing them.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education
National Center for Educational Communication

In addition to teachers, administrators should be involved in planning the teacher evaluation procedures. In fact, some propose that the design of an evaluation procedure should begin with an evaluation of how well principals and supervisors evaluate teachers. The insights gained from this experience would form a firm basis for cooperation in developing a plan for teacher evaluation.

Activities

The reason for beginning with the administrative-supervisory staff is that the activities and procedures for evaluation of professional personnel are the same, regardless of the level or position of the personnel. This does not imply that criteria will be the same, but that the methods of establishing criteria, data collection and analysis, and decisionmaking will be similar enough that the experience can be transferred from the administrative level to the teacher level.

Time Sequence

In planning the total teacher evaluation, consideration should be given to when certain activities should occur. Some experience can be gained with the administrator evaluation plan, but the time when certain phases of the processes will occur will be different. For example, the goal-setting stage or the final decision stages may occur in similar sequence to administrator evaluation but at somewhat different times of the year than with teachers. Other than legal constraints, there appears to be no reason why considerable variation could not occur from district to district or from school to school.

In addition to the sequence of events in the evaluation process, a choice must be made regarding whether the planning will occur on a piecemeal or comprehensive basis. That is, will certain aspects of the evaluation process be considered independently of other parts and decisions made to change a given part, or will the entire teacher evaluation procedure be examined and nothing changed or implemented until plans are complete for the entire process? The following comments seem appropriate to this problem area:

- The comprehensive approach is more likely to yield a well-integrated plan which satisfies the goals desired; however, it requires a mature staff and a spirit of cooperation and optimism regarding the approach.

- The piecemeal approach risks early implementation of procedures which may not be compatible with desirable activities at a later point; however, it is often argued that any steps leading to professional growth of teachers will improve learning conditions for children.

⇒ CONCLUSION

Planning teacher evaluation procedures consists of determining: (a) objectives, (b) people who will be involved, (c) activities that should occur, and (d) time sequence of events. The way decisions are made regarding this planning predetermine much of the effectiveness of the eventual procedures as well as the satisfaction with these procedures.

No. 21-C

PURPOSES OF TEACHER EVALUATION

One of the first steps in establishing or revising a program of teacher evaluation is to determine the purposes of the program. If the program is to be successful, these purposes must be identified, discussed, and agreed upon by all who are involved in the process.

Identify All Purposes

All potential purposes should be discussed openly and thoroughly; certain purposes should not be ignored just because there are critical issues involved in them or because their accomplishment may precipitate conflict. For example, one purpose for a program of teacher evaluation might be to provide a basis for sound administrative decisions regarding reemployment. Some people think that the accomplishment of this purpose creates an atmosphere which decreases cooperative relationships between administrators; they prefer to ignore this purpose and hope that the problems inherent in reemployment will not arise. However, since even in small districts these problems do arise, all purposes should be discussed openly and clarified in writing so that teachers and administrators may understand the bases for the evaluation program.

⇨ NOTE

Morale cannot be high if staff members are fearful or hostile. Since unknowns contribute to fear, evaluation programs which have written statements of purpose that are clear, precise, and complete are more likely to produce a sound basis for open communication and cooperative relationships than programs designed around ambiguous or unwritten purposes.

Some Purposes Identified

Purposes of teacher evaluation vary somewhat from district to district. They may include the following:

- To improve teaching, including out-of-classroom activities as well as classroom instruction. (This purpose is not limited to teacher behavior but implies any actions taken to improve teaching systems, the teaching environment, or teacher behavior.)

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education
National Center for Educational Communication

- To reward superior performance

- To supply information for modification of assignments (including placement in another position, reduction of load, promotion to a leadership position, or termination of employment)

- To protect individuals or the school system in legal matters (including both the protection of teachers against a capricious new administrator and the protection of the school district and children against a harmful teacher)

- To validate the selection process

- To provide a basis for career planning and individual growth and development of the teacher (including professional degrees and inservice training programs)

Each of these purposes is discussed more fully below.

Improve Instruction--There is general agreement among educators that the most important purpose for evaluating teaching is the improvement of instruction ("Teacher Evaluative Procedures," *Oregon Education*, 1966; Heald and Moore, 1968); however, this improvement may take several forms: e.g., supervisors can provide feedback regarding behavior to teachers, physical environment and materials can be modified, self-evaluation can be used to improve diagnostic skills of teachers, or information can be gathered by other teachers and discussed with the teacher.

Since evaluation of instruction is required before systematic improvement can occur, a starting point from which to work should be established. As Heald and Moore (1968:189) state, "The routes to a particular end vary according to the point of origin, and it should be one purpose of evaluation programs to establish these points." Through this knowledge of strengths and weaknesses a teacher can improve his work. Usually, when a teacher views evaluation as a means to improve his instruction, he accepts it as a part of the teaching assignment.

Reward Superior Performance--Another purpose for evaluation is to make it the basis for rewarding superior performance. However, this use of teacher evaluation tends to meet with considerable opposition from teachers, despite the fact that people outside of schools are asking why teachers should *not* be paid according to the excellence of their performance, e.g., how well pupils learn (Fishman, "Teacher Evaluative Procedures," *Oregon Education*, 1966). These increasing pressures from school boards and taxpayers for rewarding superior performance are in direct conflict with the viewpoint of the majority of teachers.

Some authors--Anthony (1968), Simon and Boyer (1967), and Howsam (1963)--suggest the teachers' major objection toward this purpose is due to the subjective nature of the evaluations. They suggest the use of objectively obtained measurements of specific behavior which have been related by research to the accomplishment of specific pupil outcomes. Teachers also resent being classified into general categories of excellence, since excellence is specific to a situation as well as a person.

Modify Assignments--Still another purpose of evaluation is to gather information for the modification of teachers' assignments, either by promotion, reduction or increase of load, or release (Heald and Moore, 1968). While these are necessary activities among all organizations, morale tends to suffer when evaluation *emphasizes* elimination of the weak and ineffective; consequently, this negative *emphasis* needs to be avoided.

⇨ CONCLUSION

The general contention is that better staff morale and a better instructional program will result from adequate and creative supervision and orderly dismissal procedures for incompetent teachers. (viewpoint: Eastmond, 1959)

Protect Individual and Organization--When the purposes of evaluation are considered from a legal standpoint, protection of individuals as well as the school organization becomes important. Responsibility for the operation of school programs is delegated by the State to the local district and from there to the teacher; with this responsibility comes the expectation that goals will be accomplished. Society does not expect schools to be operated without the same continuous evaluation that occurs in other organizations--although, from a legal standpoint, school boards have the right to establish the kind of school system they want as long as it remains within constitutional limits (Howsam, 1963; Heald and Moore, 1968). The boards' prerogatives include establishing any form of evaluation it desires. Because evaluation of teaching within a district serves as the district's protection when it is held accountable for the system it has established, evaluation is essential for legal reasons--if for no others.

⇨ NOTE

The emphasis on the legal aspects of teacher evaluation can be viewed negatively by teachers unless they realize that their own protection against unjust charges is also assured. (Smith and Tyler, 1942)

Validate the Selection Process--Though little recognized and practiced, another reason for evaluating teachers is to validate and improve the selection process. Development of procedures that link information analyzed during the selection process with teaching performance should be a part of every district's planning. The criteria used to select teachers should be consistent with those used to evaluate teachers.

⇨ SUGGESTION

Since there is a need for all who are involved in teacher evaluation to understand the relationships of evaluation to selection processes, school districts should consider involving supervisors in both selection and evaluation as one means to help them understand how the two are related. (practice: schools)

Where principals are included in the selection process, as they are in many of the schools, it is expected that they are looking for the same qualities that they will later use for evaluation. As expressed by one personnel director, "They are much more careful whom they select when they think the teacher may be placed in their own building." Also the principal may select according to his own particular needs or program, such as ungraded classrooms or "team" members with particular skills.

Promote Individual Growth--Closely related to the improvement of instruction is the teacher's own purpose for evaluation, namely, to promote self-improvement. Since most evaluation procedures assume that each teacher is interested in improving his teaching, evaluation can become the basis of planning with the teacher for individual growth and development (Howsam, 1963). To this end, new procedures begin with a goal-setting conference between the teacher and principal, and means are then established whereby the teacher can acquire and diagnose information regarding his teaching. Although the criteria are individualized in this process and the diagnosis of behavior is done by the teacher, there is usually an attempt to mesh the individual goals with the institutional goals during the first conference.

In addition, one of the functions of external evaluation is to facilitate self-evaluation. The value of self-evaluation is to allow continuous diagnosis teaching, a necessary component of good teaching. Since current ratios of principals (or other supervisors) to teachers do not allow for frequent classroom observations and conferences, no teacher can afford to rely entirely on external evaluation. All of these purposes might be expressed by saying: *The purpose of teacher evaluation is to safeguard and improve the quality of instruction received by students.* Implied in this statement is the view that evaluation of teachers will facilitate both teacher and administrator decisions (the teacher's decisions regarding how to perform as a teacher and what will improve that performance, and the administrator's decisions regarding what actions he might take to aid students and teachers).

The nature of the listing of purposes, or even of the presence or absence of a formal evaluation plan, should not alter the individual responsibility that teachers share with other professional practitioners for continuous self-improvement. However, if the school district has a well-organized program of formal evaluation with clearly specified purposes, it should be of assistance to the teacher in his voluntary program of continued self-assessment and improvement.

Changing Research View

From a research standpoint, the purposes of evaluation of teacher effectiveness have changed somewhat in recent years. For a long time, the purpose was to show that certain types of teachers provide certain patterns of teaching. In order to do this, it was necessary to observe the teacher many times to determine *what* his stable patterns of teaching were.

More recently, the emphasis has shifted in the direction of attempting to determine *why* teachers vary their behavior from one teaching episode to the next

and whether or not this variation is itself associated with particular educational outcomes. The concern here is for *variation* of teaching behavior between visits and the consequences of this variation.

For practitioners, the implication of this shift in emphasis is significant. The practitioner is interested in what works, and he realizes that some stable patterns of teacher behavior do not produce results in certain circumstances. He needs to know what produces learning in particular situations and how teachers can be encouraged to choose those behaviors which have a high probability of success.

Relation to District Goals

The teacher evaluation program should not ignore other aspects of the school program. The function of evaluation is to facilitate the accomplishment of the goals of the organization; therefore, the purposes of evaluation should be established *following* a complete review of the goals of the school district. When this has been done, the purposes of evaluation are more likely to be compatible with and contribute to the school district goals.

For example, to reward superior performance can be a legitimate purpose of evaluation, but it should be *clearly related* to the goals of the organization if it is to be included in the purposes of evaluation. If it is clear that rewarding superior performance provides incentive and motivation for creative teacher behavior which causes children to learn better, then to reward superior performance seems a reasonable purpose for evaluation.

However, some may desire to initiate a reward system for the purpose of controlling the behavior of teachers (without regard for the effect on accomplishment of educational goals), and this may lead to a misuse of the evaluation function.

⇨ CONCLUSION

The purposes of teacher evaluation programs should grow out of clearly stated goals of the school system and should contribute to the accomplishment of these goals.

IPRIP

No. 21-D

POSSIBLE PROBLEMS IN TEACHER EVALUATION

In addition to the potential benefits of evaluation, there are possible adverse effects as well. Adverse effects are indications of problems which exist, and their examination may indicate the source of a problem. Some of the possible adverse effects of teacher evaluation include:

- Human relations tensions when poor performance is perceived by the evaluator
- Reduction of creativity if evaluation system is too rigid
- Reduction of assistance provided by supervisor or principal
- Poor validation of selection process

Human Relations Tensions

There is the natural strain in human relations that results when one person evaluates another (Gruenfeld, 1966). Evaluators need to be aware of the fact that diagnosis of the teaching act *is* primarily *cognitive*, i.e., diagnosis primarily involves knowledge. However, interpretation of this knowledge by the person being evaluated is not necessarily cognitive at all, since it may be interpreted on an emotional level. When evaluation procedures include placing people in categories, e.g., "good" or "average," an emotional response is quite likely to be precipitated on the part of some teachers. Regardless of the evaluation program, evaluators should be prepared for emotional responses and should provide psychological support for the teachers being evaluated.

Reduction of Creativity

Creative teachers can function in a variety of situations. However, teachers do tend to be affected by the form of evaluation used. If the evaluation criteria require strict adherence to predetermined procedures and materials, the constraints are likely to reduce the motivation for divergence and creativity.

There is the tendency for the teacher to be shaped by a rating scale (or other measurement device) regardless of whether or not the scale validly measures good teaching. He may conform even though the measurement does not include necessary behaviors or includes behaviors not pertinent to the work. Especially

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

under situations of merit pay, states Link (1966), "A rating scale becomes a shaping device no matter how supportive the supervisor, the principal, or the system."

Reduction of Assistance

When the principal or supervisor has both the responsibility of helping the teacher improve and the job of being accountable to the school system for results, he sometimes has the tendency to emphasize accountability to the extent that it adversely affects his ability to help the teacher. At other times, his ability to assist the teacher is reduced simply because he *has* the responsibility of accountability. Some consider this reduction to be a result of a natural distrust which teachers have of the use to be made of information obtained through the evalution process. However, it may be due more to lack of knowledge of how evaluation *is* used and *must be* used by personnel who are charged with responsibilities to help teachers, viz., the evaluation information forms a basis for the counsel and assistance provided. Without it the assistance would be shallow and meaningless. The issue is not whether those who supervise and provide assistance should evaluate teachers; rather, it is how information obtained through the evaluation process should be used.

Selection Validation

A number of reasons exist for poor validation of selection procedures. However, foremost among them is the fact that very few school districts have reliable measures of teacher performance. Suppose, for example, that a principal is concerned with the fact that teachers will not accept him as a person who can help them if he emphasizes accountability to the system and attempts to measure their behavior and performance accurately. What is he likely to do and how will this affect the validation of selection procedures?

Under these circumstances, a principal is likely to evaluate all teachers more favorably than he should, indicating more uniformity than exists. In effect, his evaluation reflects a rather general "halo effect" and the measures approach a constant. When this occurs, *no predictor* can be found that will predict teacher effectiveness for that situation. Since all teachers appear to look alike once they are evaluated on the job, any predictor variable will predict as effectively (or more precisely, as *in*effectively) as any other.

The same consequences result if an evaluator allows his measurements to migrate toward a central tendency. Likewise, if the measurements provided by an evaluator are in fact not related to the behavior of the teacher or the results of this behavior, then the correlation with any predictor variable is likely to be zero. In effect, any evaluations which approach some constant value or which approach some random assignment are likely to be worthless for validation of selection procedures.

Several researchers have found that the extent of the benefit a teacher receives from evaluation is related to his attitude toward evaluation, i.e., a teacher with an unfavorable attitude benefits less than one who views evaluation positively. This implies that if teachers view evaluation as the

attempt to eliminate the ineffective rather than to help all teachers toward better and better performance, adverse effects could outweigh the benefits of evaluation ("Teacher Evaluative Procedures," *Oregon Education*, 1966).

Measurement Problems

Many of the problems underlying the adverse effects of evaluation are directly traceable to poor measurement procedures. Problems of measurement include the following:

- Prejudice, bias, or poor judgment
- Inconsistency of reaction to behavior
- Subjective ratings and classifications (or classification schemes which require high inferences to be made)
- Influence of the personality of the teacher outside the classroom on measurement of behavior in the classroom
- Attempt to measure too many elements of classroom situations (especially at the beginning of a program)
- Tendency to continue a prior viewpoint of a person's performance
- Consistent overvaluation or undervaluation

The last problem, viz., overvaluation or undervaluation, deserves special consideration. Overvaluation sometimes is caused by a desire to avoid antagonisms with teachers, an effort to make the supervisor look good, or a sympathetic response to a certain individual's problems (e.g., age or problems outside the school situation). The desire to avoid antagonisms is a strong motivator for some principals and supervisors, especially those who see themselves as placebound, desiring to live harmoniously with teachers over a long period of time, and being fearful that accurate reporting of measurements will damage the harmonious relationship.

However, the practice of indicating to a teacher that he is performing well when in fact he is not (i.e., of continually overvaluating his performance) can be as much a disadvantage to him as it is to the organization. The reason for this is that the teacher then has a limited basis for improvement in comparison with what he would have if full information were provided. Of course, the school organization and the children do not benefit either, since the performance does not improve as much as it might.

Situational Constraints

Sometimes it is difficult to determine whether poor results are due to poor performance on the part of the teacher or to situational constraints that prevent better teacher performance. When there is any doubt on this question, the conditions under which the teacher works should receive direct attention in making an evaluation of the teacher's effectiveness. Sometimes the environment (including supplies and materials as well as the psychological environment) is such that learning is inhibited, and a review of those situational factors that might influence the effectiveness of the teacher will help to prevent poor judgments regarding the teacher.

Feasibility of Annual Evaluation

Another problem of teacher evaluation is the feasibility of evaluating all teachers annually. Many school districts do not have an adequate administrative and supervisory staff to do an adequate job of evaluation. In analyzing this problem in local school districts, an attempt should be made to relate the problem to purposes for which the evaluation program is designed. However, if one of these purposes is to assist teachers in improving their instruction, it becomes very difficult to justify a program which does not include continuous evaluation of every teacher.

Consider an analogy to a baseball team. It is inconceivable that any major league team would discontinue observing the production and actions of a given player after he had been on the team for 3 or 5 years. This raises the question of why batting coaches still watch hitters. It seems reasonable to conclude that they watch batters to determine whether they are still hitting the ball, under what conditions they are hitting the ball, and whether there appear to be any flaws which hinder improved hitting.

It also raises the question of the number of coaches needed, assuming that coaching (rather than umpiring) is desired. When Mickey Mantle joined the coaching staff of the New York Yankees in September 1970, he became the fifth coach in addition to the field manager to work with 25 professional ball players. How many coaches are needed for an athletic team of x players, i.e., what is the person/coach ratio? People in policymaking positions might raise the question in comparing the student/coach ratio on the high school athletic teams with the teacher/supervisor ratio on the high school staff.

⇨ NOTE

If part of the function of the evaluator is to provide assistance for teachers (i.e., to be a coach rather than an umpire), then consideration should be given to the number of evaluators needed to do an adequate job, and efforts should be made to work toward acquisition of these evaluators.

IPRIP

No. 21-E

DEVELOPING CRITERIA FOR TEACHER EVALUATION

Generally, a criterion is defined as a standard or level of attainment against which comparisons may be made. In *evaluation* of personnel, a criterion also pertains to standards dealing with behavior and results of behavior; to assess the attainment of some process or outcome objective. This section discusses the nature of criteria dealing with teacher evaluation, and presents certain ideas found in practice and the professional literature which should be considered in making decisions about the criteria to be used.

Nature of Criteria

⇨ CONCLUSION

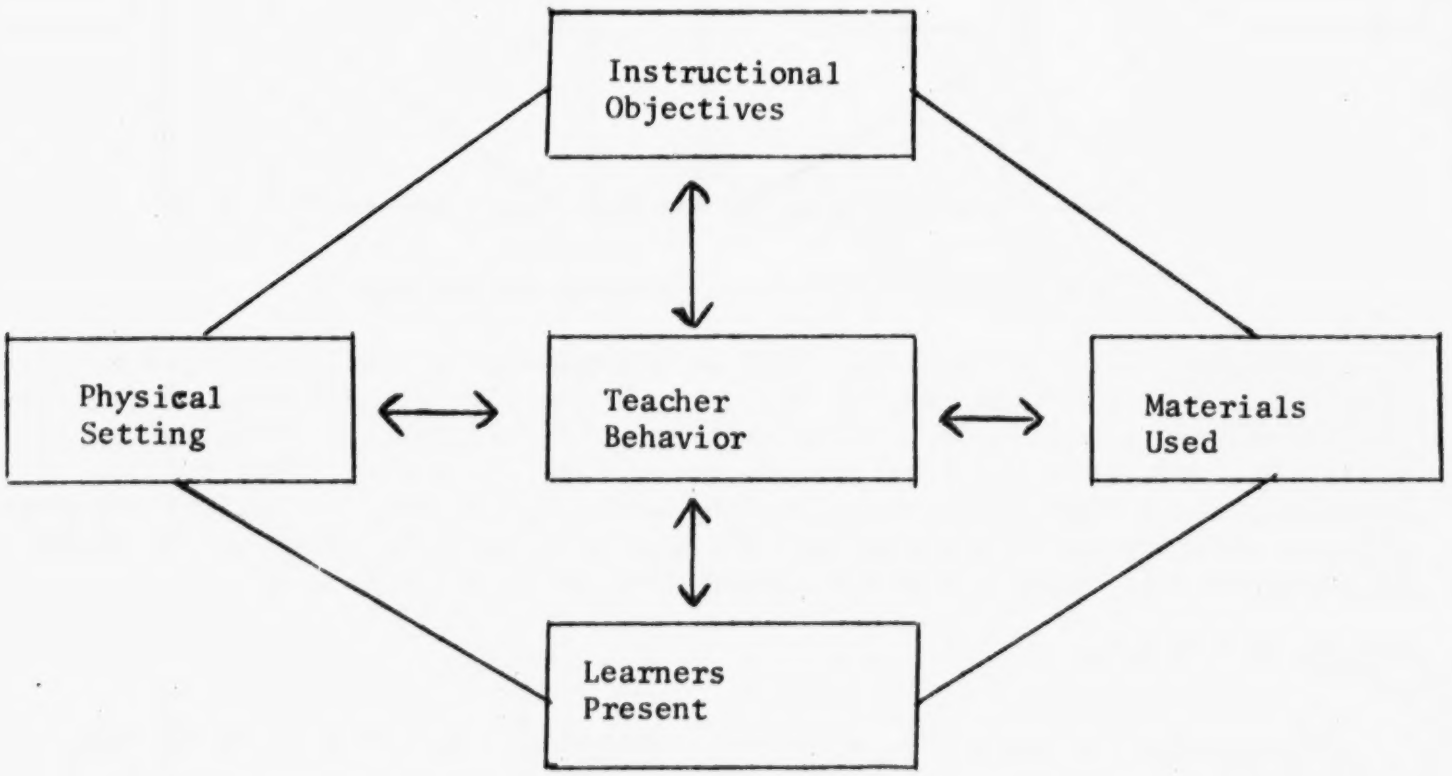
The nature of evaluation criteria is similar to the nature of selection criteria in that criteria of teacher effectiveness vary from one job to another, and they tend to change over time.
(viewpoint: Chisli, 1956; and others)

In the evaluation of teachers, most practitioners and researchers (Ware, 1964; Smith, 1967; Fishman, 1967; Ryans, 1963, 1957; Barr, 1941; Schalock, 1967) agree that a teacher functions in a highly complicated setting, in which his performance is influenced by the interaction of his personal characteristics and various situational variables. Personal variables include such components as intellectual and affective structures, perceptual habits, age, and level of training; while the situational variables might include such components as the characteristics of the learners present, the materials being used, the goals of the institution, the instructional objectives, and the characteristics of the physical setting.

Figure 1 (next page) presents representation of the interaction of the teacher's behavior and the situational components. The two-way arrows indicate the dynamic interaction between teacher behavior and the multitude of situational variables; the other lines show the interrelationships of the situational variables. As the diagram also shows, the situational variables provide the framework within which the teacher operates; however, since both teacher and environment interact with each other, there is a mediating effect on teacher performance. The implication of this diagram is that criteria determined for teacher performance should take into consideration the personal attributes the teacher brings to the job and the situational constraints placed on him by the job.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education
National Center for Educational Communication

Figure 1. A representation of the interaction of teacher behavior and various situational variables (modified from Schalock's 1967 model).



Determining Criteria

The initial step in teacher evaluation in a local school district is to determine what is considered to be important in teaching. This first step is crucial to the evaluation process, since what is considered to be important becomes the basis for developing:

- Specific teacher behaviors and results of behavior desired (identified as criteria)
- Ways to measure teacher behaviors and results, that is, criteria measurement
- Comparisons of measurements and desired outcomes (evaluation)

In determining what is considered to be important in teaching, i.e., the criteria, one should answer three questions:

- Who will decide on the criteria and their importance?
- What procedure will be used to acquire information used in making this decision?
- How will the data gathered be analyzed?

Who will Decide?—Criteria based on an individual's intuitive judgment are built on the weakest of foundations; consequently, criteria decisions would be improved if based on the *pooled* judgments of experts. Ryans (1957; 43) points out that the group of experts (jury of authorities) may consist of:

1. The *totality* of the known group of authorities or experts (e.g., all of the principals and supervisors in the school district, all members of a teachers' professional organization, all college teachers of a specified subject matter, etc.). Of course, such a procedure usually is not feasible unless the totality of experts is relatively small.

2. A *random sample* from the roster or membership list of a known group of authorities.

3. A *purposive sample* drawn from the totality of authorities as defined.

4. A *sample* of individuals who have been *specially trained* to make authoritative judgments regarding the criterion (e.g., job analysts, trained observers, etc.).

In education, method 3 probably is most often employed; however, Ryans suggests that it is the weakest of the four. He also warns that methods 1, 2, and 4 do "not necessarily insure valid criterion description, but they represent distinct improvements" (1957:43).

⇨ CAUTION

Inadequate statements describing what is important in teacher behavior often result from using biased judges. (viewpoint: Ryans, 1957:43)

A jury or group of experts in a school district, selected for purposes of criterion determination, also may include combinations of teacher, principals, supervisors, students, board members, and parents working together. Such a cooperative effort often has a positive effect on morale.

⇨ CONCLUSION

Involving teachers as well as other members of the educational community in the development of criteria may help establish more accurately defined criteria and may improve the morale of the professional staff. (practice: schools)

Where school districts require the establishment of goals for individuals, decisions regarding who will decide on what these goals will be are made by: (a) the individual teacher, (b) the principal or the supervisor, (c) the individual and the principal cooperatively, or (d) a committee composed of peers and/or administrators (the latter possibility might also include the individual teacher). Also, in establishing the broad goals for individuals, board members, students, and parents may be consulted.

⇨ CONCLUSION

Regardless of how the goals are established, output and procedural goals are more likely to be understood and attained when they are cooperatively developed by the teacher and principal (or supervisor) and are written in discriminating behavior terms. (practice: schools, industry)

What Procedures to Use?--Numerous procedures have been employed for acquiring information from people in order to determine what is considered to be important in teaching (Ryans, 1957; Chiselli, 1955; and others). Ryans (1957:44) has listed six possible techniques:

- Free-response--statements of what is important and the degree of importance, based upon the general impressions held by various members of the educational community.
- Checklist response--individuals indicate what is important and the degree of importance, on a previously compiled list of desired behaviors and outcomes.
- Position analysis--detailed systematic description of what is important for success and the degree of importance by individuals trained in carrying out such an analysis.
- Critical incidents description--detailed descriptions of actual incidents and behavior that have been observed by experts to be "critical" in learner growth and development. (Note: This technique primarily deals with teacher behavior as opposed to learner outcomes.)
- Time sampling--detailed tabulation of teacher behaviors based upon systematic observation and recording, with special attention to the conduct of observation during representative samples in time.
- Psychophysical methods--members of the jury determine what is important and the degree of importance using such methods as ranking, paired comparisons, etc.

In education, the first two techniques probably have been used more frequently than the others. However, from an objective and rational viewpoint, the first two methods are the weakest. In view of the nature of criteria for teacher evaluation, probably the method which holds the most promise is the *position analysis* technique. Evidence of the worth of this technique in improving the selection process has been provided by Palmer (1970) in his research using a completed "position analysis outline" (PAO) for improved selection decisions.

How Will Responses Be Analyzed?--Once experts respond to the criteria dimension they think important, a final judgment must be made concerning what criteria will be employed and how the criteria for evaluation will be operationally defined. Regardless of whether an individual or a group makes these decisions, a value system will be involved; however, to make such judgments without a broader perspective of others' views promotes a biased, unsystematic, and subjective approach which may tarnish the criteria descriptions.

⇨ CONCLUSION

When decisions regarding the development of criteria are based upon empirically supported and rational considerations, relevance and usability are more likely to be ensured. (position: Ryans, 1957)

A number of authors (Guion, 1961; Ryans, 1957; and Brogden and Taylor, 1950) strongly recommend that a systematic and comprehensive approach be used to develop criteria. Such an approach would be designed to provide a rational analysis of the relevancy of the possible criteria; it would hypothesize descriptions of the criterion elements and use various statistical techniques (e.g., factor analysis) to identify significant operational behaviors pertinent to the attainment of the instructional objective (Ryans, 1957).

The significance of employing a more objective approach is that evaluation of teacher behavior and learner outcomes can meet with success only to the degree that criterion judgment is based on reliable information regarding the essential attributes and behaviors involved in teaching. Consequently, such information must be based on carefully defined and rigorously controlled evaluative research and not on "armchair" methods (Turner and Fattu, 1960).

⇨ CONCLUSION

Although the relationship of teacher behaviors to student outcomes is usually limited to a specific situation, some behaviors have precipitated desirable outcomes in more than one type of situation. (research: Flanders, 1970)

Desirable Teacher Behaviors

Teacher behaviors beneficial for one group of children might not produce the same results with another. There are, however, some teacher behaviors that have precipitated desirable pupil outcomes in a variety of situations. Students seem to profit from a teacher who:

- Accepts and uses ideas and opinions of pupils
- Is flexible and adjusts behavior and strategies to situations and students
- Views teaching as a complex task which requires goal setting, individual student assessment, and decisionmaking in terms of immediate and long-range problems
- Provides students with a framework within which to interpret information

In discussing teacher behavior, one should emphasize descriptive terms prior to attaching any value to these descriptions. Likewise, avoid terms that are emotion laden, such as "democratic" or "progressive."

In addition to discussing performance (in terms of behavior and/or output) in the goal-setting conference, at least one other item should be considered: the potential of the teacher for additional tasks. Since this is directly related to the purpose of providing career counseling for teachers, it should be included in the discussion of goals to be sought.

IPRIP

No. 21-F

OBSERVING TEACHERS AND COLLECTING INFORMATION

The tools and techniques used to collect information for evaluating teachers must be related to purposes. Since multiple purposes usually exist, one question which arises is whether the system for collecting data adequately serves all of the purposes. For example, will the techniques and procedures used to accomplish the purpose of improvement of instruction serve all other purposes?

⇨ SUGGESTION

Focus on developing a system of information collection which satisfies what is considered to be the major purpose of teacher evaluation. Then examine the system developed to determine if adjustments are needed to serve all other purposes.

In making plans to collect information for the evaluation of teachers, the following questions should be considered:

- Where will the information be acquired?
- What will the information look like?
- How will a sample of total information be acquired?
- Who will collect the information?
- How much training will be required to collect it?

Sources of Information

Three basic sources of information are available for analysis and interpretation prior to evaluating teachers:

- *In-classroom behavior of the teacher*, as perceived by students, the teacher being evaluated, other teachers, administrators or supervisors, and paraprofessionals, such as teacher aides.

- *Out-of-classroom teacher behavior*, as perceived by students, the teacher, other teachers, administrators or supervisors, teacher aides, and other personnel, such as cooks and custodians.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

- *Student accomplishment,* as measured by teacher-made tests, standardized tests (achievement, attitude, or skill performance), student self-report devices, observations of student behavior (by teacher, outside observer, parent), student products or projects.

Nature of Information

Since teacher evaluation is dependent upon measurement in gathering information, care should be taken to develop procedures and train people so that appropriate instruments are chosen and used effectively.

⇨ CAUTION

Instruments should not be chosen solely on the basis of the evaluator's familiarity with the instrument, its availability, or the fact that other districts are using it. Consideration should include:

- *Relevance to goals*
- *Acceptability by those who are involved*
- *Accessibility of information*
- *Time needed to acquire information*
- *Cost*

Need To Reduce Data--In order to evaluate teachers effectively, one must reduce the available information to a form which may be analyzed and interpreted. For example, "raw" data regarding teacher behavior might be either the behavior itself or an audio or television tape recording of classroom activities. When "reduced" it might take the form of a matrix of figures or comments by an observer. Raw information regarding student accomplishment might be the answer sheets of an achievement test, while the reduced data might be a letter grade or rank in class.

⇨ NOTE

The procedure used for reducing data from its raw form influences the final interpretation of the information collected.

Types of Measures--Regardless of whether one is concerned with teacher behavior or results of this behavior (e.g., student accomplishment), when the raw data is reduced it generally takes one of the following forms:

- Rank order
- Forced distribution
- Absolute categories
- Verbal descriptors

Rank ordering is simply the ranking of individuals in a group according to some item or characteristic. A *forced* distribution requires that a certain percentage of the item being considered be placed in each descriptive category. For example, one might require that individuals be placed in five categories according to the following ratios: 10%, 20%, 40%, 20%, 10%. *Absolute category* systems describe individual behaviors or total behavior in a classroom by placing incidents or time periods into discrete descriptive categories. For example,

Flanders' Interaction Analysis System is designed to categorize verbal interaction. *Verbal descriptors* are used by observers to express what has been perceived. The descriptors may be in sentence or phrase form.

Reducing the raw data implies that some type of measurement occurs; measurement itself implies a category or numerical system that is precise. But having a category system or a numerical system does not imply that the system is precise, since the *use* of the system determines the precision.

For example, if a person were asked to count the number of shots taken in a basketball game and record approximately where they were taken and who made them, he might be able to do this with extreme accuracy. However, if the same person were asked to determine how many times a hockey player was out of position during a game, the result might be quite *in*accurate. He might record a number that looks very precise but that may deceive.

⇨ CAUTION

A wrong number is wrong even if it looks very precise. As much care should be taken in designing the implementation and use of a measurement system as in the design of the system itself.

Systematic Measurement Procedures

Observation of Classroom Behavior--Recent developments in classroom observations techniques have increased significantly the number and type of observation guides available. Simon and Boyer (1970) describe 79 different observation schedules available for use; many have been used for research only, but are easily adapted for evaluation purposes. Some instruments measure very special aspects of classroom behavior (e.g., verbal interactions), whereas others are broader in scope.

There have been two major applications of classroom interaction analysis procedures:

- To help an individual develop and control his teaching behavior
- To discover how to explain the chain of events which occur in the classroom

It seems obvious that either of these applications could be of use in evaluating teachers.

Although many guides are available and could be useful in evaluating teachers, their use is not as widespread as less precise measures. The reasons for this lack of use appear to be:

- Information has not been broadly disseminated
- Very few opportunities have been provided for learning the skills necessary to use the techniques

There are some notable exceptions to the limited use of systematic observation procedures. For example, the use of Flanders' Interaction Analysis System has increased dramatically in the Northwest because of workshops conducted by the Northwest Regional Educational Laboratory to disseminate information and develop

skills in the use of the system. Teachers and administrators find they are able to discuss teaching in a more precise language after learning the interaction analysis system.

⇨ SUGGESTION

When samples of behavior are gathered by observational means, teachers and students should be given some time to become acclimated to the presence of the observer. (viewpoint: Schalock, 1967a)

When observations are made either by a live observer or with the help of a video tape recording device, trial observations should be conducted so that teachers and students become accustomed to having an observer in the class. Research findings are somewhat sparse as to the influence observation has on those being observed.

Schalock (1967a) reports that some research (Thompson, 1963; Paul, 1944) on mother-child interaction in the home suggests that as much as 5-7 hours of observation are needed before observer influence becomes constant. Perhaps a shorter trial period would suffice in a classroom; however, until sufficient research has been conducted on this problem, data gathered using observational techniques should be interpreted with some caution.

Rating scales and checklists[2/] are used much more commonly for measuring classroom behavior than are systematic observation procedures. The major advantage of rating scales and checklists is that they allow the observer to consider clues from a variety of sources before making a judgment. However, this same characteristic can also be a disadvantage, since a delay in recording information can cause errors.

Two additional problems of rating scales or checklists should be mentioned:

- When too many ratings are clustered at a particular point, the inference is that raters are overly lenient, too harsh, or are unwilling to be decisive and objective.

- It is easier to identify the very poor or the very good than it is to differentiate in the middle range of a rating scale. Therefore, decisions concerning middle-range ratings are more difficult to justify.

2/ Probably the best single source for describing the types of analytical and general observation procedures being used in public schools is *Evaluating Teaching Performance* (1969) by the Educational Research Service of the NEA.

⇨ SUGGESTION

If rating scales or checklists are used in evaluating teachers, their accuracy may be improved by:

- *Clearly defining the focus of the evaluation*
- *Developing specific, low-inference items*
- *Using a common record form*
- *Providing adequate training for observers*

Measuring Out-of-Classroom Behavior--To the extent that the activities of the teacher in roles other than classroom instruction (such as activity in organizations, extra-class activities, contribution to curriculum development, interactions with parents, and hobbies) are considered to be important as a part of the assigned or expected responsibilities, documentary evidence should be assembled regarding these activities. As with classroom behaviors, the expectations of the school organization should be discussed with the teacher, and he should have an opportunity to set goals that are compatible with his own interests and ability to contribute.

⇨ NOTE

Since out-of-classroom activities are sometimes sources of conflict (or at least subject to differences of interpretation) between teachers and the community, new teachers should be oriented to local customs and expectations.

In measuring out-of-classroom behavior of teachers, advantages and disadvantages exist. The advantages seem to be:

- Information to be obtained is ample, often pertinent to the job expectations
- Information is relatively easy to obtain

The disadvantages appear to be:

- Reliability of information is difficult to check
- Sampling information is difficult and limited
- Information obtained is difficult to score or interpret

⇨ SUGGESTION

Written records of out-of-classroom behavior should be kept and discussed regularly with teachers in order to check the reliability and meaning of the information

Measuring Student Accomplishment--Measures of pupil outcomes include how pupils think, perform, and feel.

Traditional measures include:

- Knowledge and ability measures (what a person knows)
- Skill performance measures (what a person can do)
- Attitudinal measures (what a person feels or desires)
- Interest measures

Schools are established to facilitate pupil learning; therefore, the ultimate criterion for teacher success is the amount of learning that occurs in pupils for whom he is responsible. Thus, one of the advantages of measuring pupil growth is that it is a direct measure of outcomes that are desired. However, there are a number of disadvantages.

For one thing, there is immediate growth and there is long-term growth. The behavior of a teacher may not contribute much to the immediate learning of a pupil but may have a long-range effect on attitudes and behavior. Or conversely, short-term academic gains may be obtained at the expense of long-term negative attitudes which prevent later learning.

Another disadvantage is the difficulty of adequately controlling the situation so that growth can be attributed to the behavior of a given teacher rather than to a wide range of uncontrolled conditions which impinge on learning. Because of this problem, consideration should be given to the effect of a succession of teachers on a pupil's achievement. This should not be done *in lieu* of examining the student achievement obtained as a result of individual teachers, but should be in addition to this analysis.

Historically, student accomplishment has been avoided as a means for evaluating teachers, primarily because of the difficulties involved. However, recent emphases on accountability of teachers for productivity have caused renewed interest in setting specific student accomplishment goals and attempting to attain them. If the trend continues, it is quite likely to have a significant impact on how teachers are evaluated.

Self-Evaluation--Teacher self-evaluation may be concerned with either classroom behavior, out-of-classroom behavior, or student accomplishment. It differs from other evaluation of teacher effectiveness in that:

- There is no need for involving an external observer in the measurement process.
- Behavioral criteria are determined by the teacher rather than by some external source.

The advantage of self-evaluation seems clear: the teacher has the opportunity for improvement without external threat. The primary disadvantage is that the standards used for evaluation may not relate readily to outside criteria or needs of the school district.

Before implementing a teacher self-evaluation program, a school district should provide teachers with:

- Training to help them specify their own goals in measurement terms
- A framework (e.g., an observational system) for analyzing and interpreting their own behavior
- Technical competence needed for operating the various new media used for recording their own behavior

⇨ NOTE

Self-evaluation reduces the threat of outside intervention and, therefore, has potential for increasing motivation and creativity. However, it may not serve all purposes of evaluation, since external standards may be ignored and administrative decisions may not be facilitated.

Sampling Information

Sampling procedures employed in evaluating teachers have been designed to acquire only portions of the total amount of data available from the identified sources. In teacher evaluation, sampling techniques have been employed for two primary reasons: it is realistically impossible to collect and analyze *all* the available data (e.g., one cannot hope to observe and analyze everything a teacher does), and the techniques allow the evaluator to distribute the information demands over the available time and data sources, in order not to overtax the endurance of any individual at any given point in time (Birnbaum, 1970).

⇨ SUGGESTION

Observation of teachers should be carefully spaced over time to yield the best appraisal results. (viewpoint: Mitzel, 1967, and others)

How Often To Observe--How often should a principal observe a given teacher? A quick and obvious response is "More than most principals observe teachers," 3/ but this gives very little assistance to one seeking guidelines. A more beneficial response to this question depends on a number of factors, including the purposes for the observation and the resources available. To satisfy most purposes, each teacher should be observed in the classroom environment several times annually at different times of the day or in varying types of instruction. Such observations can be a mixture of observation procedures and may be made by appointment or not.

- Specify the amount of time per week which should be spent in teacher evaluation procedures.
- Indicate how long each observation should be and whether the observation should be followed by a teacher conference and written record.

3/ In an NEA survey (1969), 80 percent of the responding schools indicate that they evaluate probationary teachers more often than tenure teachers; yet, 85 percent of the probationary teachers were evaluated no more than twice annually.

- Determine the approximate number of teachers who can be observed annually.
- Determine how many times each teacher should be observed annually.

To see what would happen if the above procedure were followed, create a hypothetical situation in which:

- The principal should spend 5-8 hours per week in teacher evaluation procedures.
- Length of classroom observation is approximately 30 minutes
- Observation will always be followed by a teacher conference.
- Written records will be kept of observations and conferences.
- There are 60 teachers in the school.

For this example, the principal would be able to observe approximately five teachers weekly, on the average; this would allow him to make a systematic observation of each teacher approximately every 12 weeks or only three times annually.

Is this sufficient for accomplishing the purposes established for teacher evaluation in this district? If not, then steps should be taken to either:

- Establish more reasonable goals for the evaluation program, or
- Modify the procedures established (e.g., spend more time in observation weekly), or
- Acquire more resources for doing the task

If more assistance is needed, then the principal and the personnel in the central office should agree on who and what the nature of the assistance should be. For example, should the vice principal, department heads, consultants, and central office personnel be involved? If so, should their role in the total evaluation process be different from that of the principal?

⇨ NOTE

Some teacher evaluation programs are designed for failure because not enough personnel are provided to do the job adequately. Personnel inadequacies should be relatively easy to identify, and a number of alternatives may exist for remedying the problem.

Who Collects Information?--To evaluate teachers, information should be acquired regarding classroom behavior, out-of-classroom behavior, and student performance. *Who* collects this information depends on the plan which is implemented, and the following people should be considered when developing a plan:

- Principals
- Vice principals
- Department heads
- Subject matter specialists (consultants, supervisors)
- General consultants
- Personnel specialists or directors
- Peers (other teachers)
- Students
- Parents

It seems reasonable to expect that different people would collect different types of information. For example, a principal of a large high school would probably collect different types of information from that of the department head or the subject matter specialist. Because of the difficulty of having expert knowledge in all of the subject matter specialties of the various teachers, the principal is more likely to provide meaningful feedback to teachers in the areas of classroom atmosphere, general learning principles, and verbal and nonverbal interaction.

The subject matter specialist or department head, on the other hand, could collect information concerning the appropriateness of the content and activities related to the content. The feedback that he is able to provide to teachers regarding content and activities that are appropriate for the given subject specialty should be of benefit to teachers. The judgments made about these aspects of teaching are part of the specialty of the consultant or supervisor who should be made to feel a part of the process of teacher evaluation as much as other people who collect and analyze information regarding the teacher.

Difficulties arise with regard to *use* of information collected by subject matter specialists. Some seem to fear being labeled a "tattler" or to feel that the task of collecting information for evaluation purposes is outside their role. A clear specification of the role of the subject matter specialist and of the use of the information he collects should help to remove some of these difficulties. This specification is the job of the administrator, who has a responsibility to set objectives that help to harmonize the activities of people with goals of the organization.

Training Data Collector

A final consideration in planning for data acquisition concerns the training of the data collector. If the data to be collected are to be in raw form, little training of the collector is required. However, if data are to be reduced at the time of collection, then the collector must be capable of accurately recording and reducing the data for analysis and interpretation.

For example, when studying the interpersonal communications between a teacher and students, an observer (data collector) must provide data which are valid and reliable. Therefore, the observer must undergo extensive training to memorize the definitions of the behavior categories used for classifying teacher and student messages; he must master the use of the behavior record form and the other recording procedures so that accurate judgments can be made; and he must demonstrate that his records are consistent from one observation to another.

⇨ SUGGESTION

Individuals responsible for acquiring "reduced" data should be provided extensive training in reduction procedures to ensure valid and reliable information. (viewpoint: research staff)

Prerequisites to Good Observation

Most information is collected by means of observation, and good observation has certain prerequisites:

- Some purpose needs to be identified. A person does not just observe; he observes *for* something. He does not just look; he looks for something specific.
- The more specifically one identifies what he is looking for, and the more systematically he plans for observation, the more likely it is that he will know something following the observation.
- What is observed needs to be subject to checks and controls, in order that some determination can be made of the validity, reliability, and precision of the observation.

Limits to Observation

Observation is direct. It is not an indirect predictor of behavior as a test is, but rather tells something directly about the behavior of teachers. As a result of this directness, it has some limits. For example, some events cannot be predicted and, therefore, are difficult to observe. A teacher who has a very sensitive ability to handle delicate human relations problems in the classroom may not be able to demonstrate this for an observer because of the difficulty of predicting when an appropriate situation will occur. Likewise, the duration of events is a very practical limitation to observation of classroom events. It sometimes is hard for an observer to see the continuity and sequences of events that make some teachers extremely effective.

Once the prerequisites to good observation have been provided, what about the act of observing? The observer's mental set during observation is quite important. Otherwise, he cannot interpret gestures, expressions, etc. This means that he must know something about the context within which he is observing, and implies that he should:

- Discuss the situation which he will observe with the teacher prior to observing
- Confer with the teacher following the observation to check his own understanding of the context
- Develop his own understanding of the impact of contexts on both students and teachers

An observer's introspection and experience can be both a hindrance and a help. It can cause him to overlook and misinterpret as well as to be very shrewd in perceiving subtle differences.

⇨ SUGGESTION

Observers of teacher behavior and classroom interaction should develop means for checking their own reliability; this usually means comparing observations with another observer and the teacher being observed.

Because an observer is usually responsible for knowing something about the situational context, the interaction among people, and individual behavior, he should attempt to increase reliability of observation by:

- Adequately defining what is to be observed
- Examining his own background and experience to determine whether it might be distorting his perception
- Establishing categories which assist in recording behavior
- Comparing observations with others to help establish and maintain reliability

Guidelines

Some guidelines for development of information collection procedures for use in evaluating teachers include:

- School board policy should identify all purposes of evaluation, specify general procedures to be used in evaluation of personnel, and identify what types of data the school board desires in order to provide information to the public and to set policy.

- All new teachers should be oriented to the total procedures used, as well as the forms and reports that will be used.

- Vice principals and department chairmen should be involved in evaluation procedures, and all persons involved should be acquainted with their role.

- Student reactions to teacher behavior and to classroom activities should be encouraged on an anonymous basis at the option of the teacher.

- If peer evaluation is used, training in making honest and helpful comments should be provided.

PREP

No. 21-G

MEASURING INSTRUMENTS

In the evaluation of teachers, measurement techniques are central to the acquisition of data from the standpoint that:

- They sometimes influence the kind of information to be gathered, i.e., in considering *what* data are to be gathered, it is necessary to contemplate the *type of measure* that will be used
- They help order the available data
- They help reduce the error of informal human observation

As Schalock (1969, V-26) so aptly put it:

> *In the absence of instruments for the extension of the senses or for the control of conditions, human observations are liable to error. Instruments are a means for approximating more closely the property under observation.*

⇨ CONCLUSION

Teacher evaluation is dependent upon measurement as a basis for information gathering, because it is through measurement that the evaluator ascertains the quantity or quality of something. (viewpoint: Schalock, 1968)

Selection of Instruments

In selecting measures for evaluation, a major rule of thumb is "select the instrument which best fits your purpose," i.e., identify the measurement techniques and strategies which provide the data desired. Although this guideline is quite simple, too often choices are made on the basis of familiarity with or easy availability of instruments rather than because of their appropriateness.

To help in making appropriate decisions regarding the kinds of measures to use or develop, Lyons (1970) has suggested four practical considerations or resistants:

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education
National Center for Educational Communication

- *Cost Factor* - Priorities must be determined for the kinds of data needed and decisions made to allocate money among these priorities.

- *Time Factor* - Some measures take a great deal of time to use and to develop properly; and if not enough lead time is available, the use of such instruments will not be feasible.

- *Source Factor* - It does no good to decide on a particular instrument that would do the job, allocate appropriate resources, and then find out it is not possible to collect the data because no data source is available.

- *"Taboo" Factor* -An otherwise satisfactory instrument can meet with resistance if it conflicts with local traditions or custom.

Characteristics of Instruments

Besides the practical considerations listed above for choosing or developing an instrument for teacher evaluation, one should also consider the characteristics which indicate adequacy of *any* measuring instrument, viz., relevance, reliability, validity, fidelity, and ease of administration. A number of authors (Lyons, 1970; Schalock, 1968; Kerlinger, 1964; Thorndike and Hagen, 1962; Ryans, 1957; and others) describe those characteristics as follows:

- *Relevance*--This quality is sometimes referred to as validity, i.e., the extent that the instrument appears to measure what it says it does.

- *Reliability*--This quality concerns the consistency or reproducibility of the measure, i.e., the instrument continues to maintain its stability from one application to the next.

- *Validity*--This quality pertains to the fact that the instrument measures that behavior, object, or event for which it was intended to measure.

- *Fidelity*--This quality relates to the degree to which the response to the instrument parallels the true or actual performance (e.g., skill tests in physical education have a greater degree of fidelity than do paper and pencil tests).

- *Ease of Administration*--This quality involves the practicality of the instrument in the evaluation, i.e., its availability, scoring ease, etc.

⇨ CONCLUSION

Data acquired for teacher evaluation purposes may be analyzed and interpreted with a greater degree of confidence if in the evaluation process the following two questions can be answered positively:

- *Will the measuring instruments employed fit the purpose of the evaluation and will they do an adequate job?*

- *Have the measuring instruments been implemented accurately?*

(viewpoint: research staff)

IPRIP

No. 21-H

POST-OBSERVATION CONFERENCES, COMMUNICATION

As clear and precise communication is essential to the establishment of acceptable purposes for teacher evaluation, so excellent communication between teachers and evaluators is essential in the post-observation conferences and formal reports. Every classroom observation should be followed by a discussion between the teacher and the cooperating evaluator(s), and this discussion should take place as soon after the observation as possible. There should be open communication regarding what will be reported to the central office, and written copies of any report should be given to the teacher.

When To Make Post-Observation Decisions

Since one of the functions of the total evaluation process is to make decisions possible, one of the first considerations an evaluator must face is *when* he will make the decisions that are based on observations. He has three choices; he can make the decisions:

- Before the post-observation conference, allowing him to use the conference for informing the teacher of the decision
- During the conference, allowing the teacher to be a part of the decision process
- Following the conference, using the conference as an opportunity to collect additional information that might be pertinent to the decision

Some practitioners favor one of these times for all decisions, but others tend to choose one in terms of the purpose to be accomplished by the particular conference or decision. For example, if a certain decision needs to be made regarding classroom procedures and the evaluator is providing information to be of assistance to the teacher, the decision might be made during the conference. However, if a decision is to be reached regarding retention of the teacher, the evaluator may delay the decision until after the conference in order to use the information obtained from the conference (or possibly from other sources) to make final decisions.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

Post-Observation Conference

In both business and schools some kind of meeting between supervisor and employee often follows observations of the subordinates' work. Usually this meeting is referred to as an " evaluation conference" by educators and as a "performance appraisal interview" by business personnel. Essentially, both are used to fulfill the same purposes.

In business the diversity of these purposes has greatly increased since the appraisal interview has been seen as a means to *develop* subordinates as well as evaluate them (Solem, 1960).

⇨ NOTE

The performance appraisal interview can involve such diverse functions as the evaluation of performance, motivation of the subordinate, warning, praising, developing, treating the subordinate as an individual, recommending future courses of action, and the differential granting or withholding of an entire system of rewards and punishments. (research: Solem, 1962)

Research Findings

Research on the use of post-observation conferences indicates the following:

- Criticism has a negative effect on employees; it tends to build defensiveness.
- Praise has very little effect on future productivity.
- Mutual goal-setting for the future improves performance.
- Assistance and coaching effect better results when it is done daily rather than once yearly.
- Teachers accept decisions more readily if the focus is on improving performance and the situation.
- The number of improvements that can be accomplished at any one time is limited; therefore, one should choose a few and focus on them. (This probably implies the need to develop a specific strategy for assisting teachers.)

Developing a Plan for Conferences

To develop a workable plan for post-observation conferences, one should consider the following:

- Provide a written guide to aid principals and supervisors in conducting post-observation conferences.

- Provide demonstrations and conduct practice sessions.
- Discuss how the following topics and activities can be included in the post-observation conference:

  --Purpose of the interview

  --Description of favorable information (feedback should be honest rather than effusive praise)

  --Discussion of weak aspects of performance (constructive criticism must be given in friendly, cooperative spirit)

  --Asking for reactions

  --Responding, discussing (help the teacher to know how well he is doing and what is expected)

  --Considering appropriate action with teacher

  --Determining what additional information is needed, when it will be sought

  --Planning the next steps

  --Concluding the interview

Sources of Conflict

The primary problem with the followup conference is that it can evolve into a conflict situation which creates a gap between administrator and classroom teacher. This is especially true if:

- There is no preobservation conference to establish goals and study the context
- There is no agreement on roles and responsibilities
- There has been no assistance given to the teacher in planning his work
- There is no assistance given in the post-observation conference (only discouragement)
- There is no opportunity for feedback regarding how well the administrator is doing his job

If an administrator or supervisor is concerned with the possibility that a conflict situation might arise and desires to take steps that might reduce the gap between himself and a classroom teacher, then he should:

- Establish open, authentic communication with teachers
- Seek agreement between the goals of the school system and the individual aspirations of teachers
- Be willing to be evaluated by teachers on those aspects of his job which affect the teachers

Especially pertinent to the open communication between administrators and teachers is the necessity to maintain open files of formal written evaluations. When teachers know what is written regarding their performance and what is reported to the board of education, anxiety and rumor are reduced.

When the file is closed to the teacher, he assumes that the file contains confidential recommendations obtained for promotions. Under the circumstances, some systems provide copies of all evaluation reports to the teacher; some systems require four copies so that the principal, the personnel director, and the assistant superintendent (or curriculum personnel) receive copies also. One central office person is then responsible for reviewing each evaluation report to suggest actions that might be taken for improving individual teachers or to recommend changes in the system. When the official record system is open to the teacher, principals sometimes keep a personal file which is not officially a part of the district's record system.

IPRIP

No. 21-I

ASSESSMENT OF THE TEACHER EVALUATION PROCESS

Need for Assessment

Too often, after a system has been developed for the evaluation of teachers, schools neglect to develop a systematic approach to judge the effectiveness of that system. If the teacher evaluation process is not periodically analyzed for problems, and if concern for assessment of teacher evaluation only occurs when any one particular teacher seems in difficulty and is in jeopardy of losing his position because of lack of competence or effectiveness, it becomes very difficult to discover who really is at fault: the individual teacher, the system for teacher evaluation, some aspect of implementing the evaluation system, or a combination of these elements.

When a school district continuously monitors the teacher evaluation process, it has constant sources of feedback, which allows for anticipation of problems and, as in the case of the ineffective teacher, may point to modification in supervision before problems reach a point where alternatives are limited to forced resignation or cancellation of contract.

Since the plan for evaluation should be comprehensive enough to examine the entire evaluation program, the means for gathering information about the effectiveness of the evaluation process should be planned at the same time the evaluation procedures are planned.

Analysis of Problems

An analysis of the total evaluation process should include examining the realism of the goals of the process, the effectiveness of the teaching procedures, and the adequacy of implementing the procedures decided upon. In examining these aspects, one should seek answers to the following questions:

- Is the instruction improving?
- Are teachers receiving assistance?
- Are students learning?

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

- Are teachers with problems improving?
- Are consistently ineffective teachers being released?
- Do teachers understand what is expected of them?
- Is adequate information being provided to improve the selection process?
- Is the board of education provided adequate information for making personnel and policy decisions?

Some of these questions may be answered by obtaining answers from teachers or administrators, students, or parents. Others may be answered by interviewing those teachers who either resign or are released. Businesses and industry use the exit interview very effectively to acquire information regarding:

- Why individuals leave the organization
- Employee's perceptions of problems in supervision and evaluation
- Employee's perception of problems in selection and placement procedures

Some evaluation problems may not derive from the system itself nor the intended procedures but may be due to problems of *implementation*. For example, if adequate time is not spent in observation and providing feedback to teachers, the goal of instruction improvement may not be served. On the other hand, a thorough examination of *goals* may indicate that they are unreasonable for a particular teacher or group of pupils. Other sources of problems may be *procedural*.

Analysis of Information

If analysis of the data collected on the evaluation process is needed, the revision begins at the point where change should take place. For example, where data collection procedures are inadequate, then a review of *purposes* is not needed, but rather one should change the data collection *procedures*. Assessment may indicate that principals need more training in observation procedures or in methods of feedback to teachers.

Assessment of teacher evaluation is not a closed system. It occurs within the larger context of the total operation of schools and within the context of the local community and society.

⇨ CAUTION

When school districts do not consider the multiple contexts within which the teacher evaluation system operates, they increase the probability of ignoring important considerations in planning for problems that may occur. These contexts include:

- *Human relations with teachers and community*
- *Development of school policy*
- *Teacher training*
- *Teacher organizations and negotiations*
- *Evaluation of other personnel*
- *Total school program evaluation*

Human Relations with Teachers and Community--Because evaluation of teachers is a human process as much as it is an organizational program, continuous efforts must be made to develop reliable measures of effective teacher behaviors which relate to pupil outcome in specific situations. Pupils, parents, and other community members are at the very least indirectly involved in teacher evaluation; and school districts must help them understand just how and to whom teachers and administrators are accountable.

Development of School Policy--School policy, developed by local school boards, is what gives direction to administrators; it is the public expression of a philosophy. Administrators must not only adhere to personnel policies which reflect that philosophy, but professionals have an obligation to inform and interact with the public and thus provide means to policy change.

Teacher Training--Teacher evaluation seems to be increasingly more interrelated with teacher training. As new certification standards and intern-type training programs evolve, the new information about teacher effectiveness must be fed back into the personnel policy development effort of local school districts.

Teacher Organizations and Negotiations--No longer do teacher organizations divorce themselves from issues involving evaluation and accountability. Many teacher organizations have already acquired negotiation agreements with local boards. State and national organizations have obtained necessary legislation which gives teacher organizations bargaining powers.

Regardless of one's stand on the issue of teacher rights, and whether or not a labor-management relationship will emerge between teachers and administrators, negotiations will occur and they will include discussions about and planning for improved teacher evaluation processes. This fact should give even more impetus to local efforts regarding effective teaching and systems for evaluating that teaching.

Evaluation of Other Personnel--Systems for evaluating teacher effectiveness are, of course, related to the quality of those who administer that system and to the quality of those who administer that system and to the quality of those who are part of it (secretaries, cooks, etc.). No less rigorous evaluation procedures should occur, and consequently no less comprehensive personnel systems should be developed for school personnel other than teachers. Some believe that only when systematic evaluation of *evaluators* (e.g., principals, supervisors) occurs will teachers more readily accept accountability functions.

Total Program Evaluation--The evaluation of teachers is part of the larger efforts a school system makes in assessment of the total program. For example, changes in curriculum, groupings of pupils, school plant design, and instructional materials have an effect upon and are affected by teacher evaluation.

Training Evaluators

Some problems of teacher evaluation may be solved through additional training of evaluators. Even the best principals and supervisors may need to be trained to avoid allowing their personal biases and prejudices to affect the accuracy of observations. They may need to be trained in using observation procedures which use definitions of behaviors to be observed and standards to be applied to each behavior.

Public school systems treat the training of evaluators much more casually than industry does, in spite of evidence that training is likely to increase:

- Validity of decisions
- Reliability or consistency of decisions
- Discrimination of measurement
- Feelings of certainty regarding decisions

Procedures used by many businesses and industry and some school districts for improving evaluator performance include:

- Elective inservice courses
- University courses
- Group meeting devoted to evaluation
- General explanations given at regular administrative meetings
- Workshops or clinics lasting from 1 to 3 days (including assistance from outside consultant, practice, discussions, use of multimedia presentations)
- Written documents or manuals
- Individual consultation

Summary

The *evaluation of professionals* provides information needed to judge effectiveness of the individual teachers and allows better judgments to be made about modifications in training and placement. The *assessment of the evaluation process* gives the information needed to make judgments about the effectiveness of that system, including how well the system:

- Measured teacher goodness
- Planned the process
- Implemented the system
- Trained and supervised those who are evaluating teachers

Even if teacher effectiveness is measured carefully by process and product, and even if *evaluators have been trained* to observe teachers, only when the *assessment* of that *process* is precise and systematic will teacher evaluation contribute fully to the total enterprise of education.

IPRIP

No. 21-J

BIBLIOGRAPHY

Amidon, Edmund J., and John B. Hough, eds. *Interaction Analysis: Theory, Research and Application*. Reading, Mass.: Addison Wesley Publishing Co., 1967.

Anthony, Bobbie M. "A New Approach to the Merit Rating of Teachers," *Administrator's Handbook*. Chicago: University of Chicago, September 1968.

Barr, A. S. "Teaching Efficiency," *Encyclopedia of Educational Research*. New York: The MacMillan Co., 1941.

Birnbaum, Martin J. "Information Collection Procedures," *A Strategy for Evaluative Design*. Monmouth, Oreg.: Oregon State System of Higher Education, 1970.

Brogden, H. E., and E. K. Taylor. "The Theory and Classification of Criterion Bias," *Educational and Psychological Measurement*, 10:159-186, Summer 1950.

Buros, Oscar Kristen, ed. *The Sixth Mental Measurements Yearbook*. Highland Park, N. J.: Gryphon Press, 1965.

Eastmond, Jefferson N. *The Teacher and the School Administration*. Boston: Houghton Mifflin Company, 1959.

Educational Research Service. *Evaluating Teaching Performance*, Washington, D.C.: American Association of School Administrators and Research Division, NEA, No. 3, 1969.

Ellena, William J., Margaret Stevenson, and Harold V. Webb. *Who's A Good Teacher?* Washington, D.C.: National Education Association, 1961.

"Evaluation of Teaching Competence," *NEA Research Bulletin*, 47:67-75, October 1969.

Fishman, Joshua A. "Cross-Cultural Perspective on the Evaluation of Guided Behavioral Change," *The Evaluation of Teaching*, Pi Lambda Theta, 1967.

Flanders, Ned A. "Intent, Action and Feedback: A Preparation for Teaching," *The Journal of Teacher Education*, 13:251-260, September 1963.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication

_____. *Analyzing Teaching Behavior*. Reading, Mass.: Addison-Wesley Publishing Company, 1970.

Gage, N. L. "A Method for 'Improving' Teacher Behavior," *Journal of Teacher Education*, 14:261-266, September 1963.

_____, ed. "Paradigms for Research on Teaching," *Handbook of Research on Teaching*. Chicago: Rand McNally and Company, 1963.

Ghiselli, E. E. "Dimensional Problems of Criteria," *Journal of Applied Psychology*, 40:1-4, February 1956.

Gruenfeld, L. W., and Peter Weissenberg. "Supervisory Characteristics and Attitudes toward Performance Appraisals," *Personnel Psychology*, 19:143-151, Summer 1966.

Guilford, J. P. Psychometric Methods, New York: McGraw Hill Book Company, 1954.

Guion, Robert M. "Criterion Measurement and Personnel Judgments," *Personnel Psychology*, 14:141-149, Summer 1964.

Heald, James E., and Samuel A. Moore, II. *The Teacher and Administrative Relationship in School Systems*. New York: The MacMillan Company, 1968.

Hill, William Morris. "The Effects on Verbal Teaching Behavior of Learning Interaction Analysis as an In-Service Education Activity," Doctoral thesis, Ohio State University, 1966. *Dissertation Abstracts*, 27A:2084.

Hough, J. B., and R. Ober. "The Effects of Training in Interaction Analysis on the Verbal Behavior of Preservice Teachers." Paper presented at the annual meeting of the American Educational Research Association, 1966, Chicago, Illinois.

Howsam, Robert B. "Teacher Evaluation: Facts and Folklore," *The National Elementary Principal*, 43:7-18, November 1963.

Jensen, Paul J. *A Study of Self-Evaluation Applied to In-Service Education*. Bureau of Research, U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 5-1121. Monmouth, Oreg.: Teaching Research, Oregon State System of Higher Education, 1968.

Kerlinger, Fred N. *Foundations of Behavioral Research: Educational and Psychological Inquiry*. New York: Holt, Rinehart and Winston, 1964.

Kliebard, Herbert M. "The Observation of Classroom Behavior: Some Recent Research." *The Way Teaching Is*, Association for Supervision and Curriculum Development and the National Education Association Report on the Seminar on Teaching, 1966.

Link, Frances R. "Merit Rating: Have the Issues Changed?" *Supervision: Emerging Profession*. Washington, D.C.: Association for Supervision and Curriculum Development, 1969.

Lopez, Felix M. *Personnel Interviewing: Theory and Practice*. New York: McGraw Hill Book Company, 1965.

Lyons, Tom. "Instrumentation," *A Strategy for Evaluation Design*, Monmouth, Oreg.: Teaching Research, Oregon State System for Higher Education, 1970.

Malone, C. F. "Design for Productive Classroom Evaluation," *Clearing House*, 42:348-354, February 1968.

McNeil, J. D. "Concomitants of Using Behavioral Objectives in the Assessment of Teacher Effectiveness," *Journal of Experimental Education*, 36:69-74, Fall 1967.

Medley, Donald M., and Russell A. Hill. "Dimensions of Classroom Behavior Measured by Two Systems of Interaction Analysis," *Educational Leadership*, 27:821-824, May 1969.

_____ and H. E. Mitzel. "Measuring Classroom Behavior by Systematic Observation," *Handbook of Research on Teaching*, Chicago: Rand McNally, Inc., 1963.

Merlino, Edward Stephen. "A Study of Teacher Behavior and Attitudes Based on the Results of Pupil Feedback." Master's thesis, University of Washington, Seattle, 1965.

Meyer, H. H., Emanuel Kay, and J. R. P. French, Jr. "Split Roles in Performance Appraisal," *Harvard Business Review*, 43:123-129, January-February 1965.

_____ and W. B. Walker. " A Study of Factors Relating to the Effectiveness of a Performance Appraisal Program," *Personnel Psychology*, 14:291-298, Autumn 1961.

Mitzel, H. E. "Teacher Effectiveness," *Encyclopedia of Educational Research*, New York: The MacMillan Company, 1960.

_____. *Criterion Problems in Validating Teacher Selection Policies and Procedures."* Brooklyn, N.Y.: New York City Board of Education, June 1967.

Moscowitz, Gertrude. "The Attitudes and Teaching Patterns of Cooperating Teachers and Student Teachers Trained in Interaction Analysis," *Interaction Analysis: Theory, Research and Application*, Addison-Wesley Publishing Company, 1967.

Noland, Robert L., and J. J. Moyland. *How to Conduct Better Performance Appraisal Interviews*. Springdale, Conn.: Motivation, Inc., 1967.

Odiorne, George S. *Management by Objectives*. New York: Pitman Publishing Corporation, 1965.

Palmer, Dale Harlie. "Situational Variables Affecting Decision Making in the Selection of Teachers." Doctoral thesis, University of Washington, Seattle, 1970.

Parsons, Theodore W. "Guided Self-Analysis System for Professional Development Education Series." Berkeley, Calif.: University of California, 1969.

Paul, Janet S. "Observer Information of the Interaction of a Mother and a Single Child in the Home." Unpublished master's thesis, Oregon State University, Corvallis, 1964.

Popham, W. James, and Lloyd S. Standlee. "Out-of-School Activities and Professional Performance of Teachers," *Bulletin of the School of Education*, Indiana University, 34:27-29, July 1958.

Poppleton, Pamela. "The Assessment of Teaching Practices: What Criteria Do We Use?" *Education for Teaching*, 75:59-64, Spring 1968.

Redfern, George B. *How to Appraise Teaching Performance*. Columbus, Ohio: School Management Institute, Inc., 1963.

Rogers, Carl. "Towards a Theory of Creativity," *ETC: A Review of General Semantics*, 11:249-260, Summer 1954.

Rose, Gale W. "The Effects of Administrative Evaluation," *The National Elementary Principal*, 43:50-56, November 1963.

Rosenshine, Barak. "Evaluation of Instruction," *Review of Educational Research*, 40:279-300, April 1970.

Ryans, David G. "Notes on the Criterion Problems in Research, with Special Reference to the Study of Teacher Characteristics," *Journal of Genetic Psychology*, 91:33-61, September 1957.

_____. "Assessment of Teacher Behavior and Instruction: Observation and Assessment of Teaching," *Review of Educational Research*, 33:423, October 1963.

Schalock, H. Del. "Issues in the Conceptualization and Measurement of Teaching Behavior." A paper presented at the American Educational Research Association Annual Conference, Chicago, 1967a.

_____. *The Teaching Research System for the Description of Teacher Behavior in Context*. Monmouth, Oreg.: Oregon State System of Higher Education, 1967b.

_____. "Measurement," *CORD: National Research, Evaluation, and Development Training Manual*, 2d ed., Monmouth Oreg.:: Teaching Research, Oregon State System of Higher Education, 1969.

_____ and James H. Beaird. *Increasing Prediction of Teacher's Classroom Behavior through Use of Motion Picture Tests*. Cooperative Research Report No. 5-0836. Monmouth, Oreg.: Teaching Research, Oregon State System of Higher Education, July 1968.

Short, Robert Allen. "The Relationship of Teachers' Classroom Behavior to the Achievement of Junior High School Students and the Effect of Interaction Analysis Feedback on Teachers' Classroom Behavior." Doctoral thesis, University of Washington, Seattle, 1968.

Simon, Anita, and Gil Boyer, eds. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments, Vol. I--XIV*. Philadelphia, Pa.: Research for Better Schools, Inc., 1967-1970.

Smith, B. Othanel. "Teaching: Conditions of its Evaluation," *The Evaluation of Teaching*, Pi Lamda Theta. 1967.

Smith, E. R., and R. W. Tyler. "Appraising and Recording Student Progress," *Adventure in American Education.* Progressive Education Association, Commission on the Relation of Schools and Colleges, New York: Harper and Brothers, Inc., 1942.

Soar, R. S. "New Developments in Effective Teaching," *American Biology Teacher,* 30:43-47, January 1968.

Solem, Alan R. "Some Supervisory Problems in Appraisal Interviewing," *Personnel Administration,* 23:27-35, May-June 1960.

Super, Donald E., and John O. Crites. *Appraising Vocational Fitness.* New York: Harper and Row, 1962.

Taft, Ronald. "The Ability to Judge People," *Psychological Bulletin,* 52:1-23, January 1955.

"Teacher Evaluative Procedures," *Oregon Education,* 41:16-17, 30, October 1966.

Thompson, Joann N. "Observer Information of Mother with Three Pre-School Children in Home." Unpublished master's thesis, Oregon State University, Corvallis, 1963.

Thoreson, Carl E. "Video in the College Classroom: An Exploratory Study," *Personnel and Guidance Journal,* 45:144-148, October 1966.

Thorndike, Robert L., and E. Hagen. *Measurement and Evaluation in Psychology and Education.* New York: John Wiley and Sons, Inc., 1962.

Turner, R. L., and N. A. Fattu. "Skill in Teaching, A Reappraisal of the Concepts and Strategies in Teacher Effectiveness Research," *Bulletin of the School of Education, Indiana University,* 37, May 1960.

_____ and others. "Skill in Teaching, Assessed on the Criterion of Problem Solving: Three Studies," *Bulletin of the School of Education, Indiana University,* 39, January 1963.

Ware, Claude T., Jr. "Individual and Situational Variables Affecting Human Performance," *Human Factors,* 6, December 1964.

Webb, E. J., and others. *Unobtrusive Measures: Nonreactive Research in the Social Services.* Chicago: Rand McNally and Company, 1966.

Wilk, Roger E., and William H. Edson. "Predictions and Performance: An Experimental Study of Student Teachers," *Journal of Teacher Education,* 14:308-317, September 1963.

IPRIP

No. 21-K

ERIC DOCUMENTS ON TEACHER EVALUATION

Following are some of the most recent documents on evaluating teacher performance entered into the ERIC system. These documents are available in microfiche (MF) and hard copy (HC) from the ERIC Document Reproduction Service, 4936 Fairmont Boulevard, Bethesda, Maryland 20014 at the prices cited.

ED 039 860. *Effective University Teaching and Its Evaluation.* 41 pp. MF - 25¢; HC - $2.15.

ED 398 289. *Attitudes and Characteristics of Effective and Not Effective Teachers of Culturally Disadvantaged Children.* 18 pp. MF - 25¢; HC - $1.

ED 039 170. *The Development of an Instrument Designed To Secure Student Assessment of Teaching Behaviors That Correlate With Objective Measures of Student Achievement.* 29 pp. MF - 25¢; HC - $1.55.

ED 039 197. *Identifying Effective Teaching Behaviors.* 136 pp. MF - 75¢; HC - $6.90.

ED 039 200. *Research Related to the Evaluation of Graduates of Teacher Preparation Programs.* 19 pp. MF - 25¢; HC - $1.05.

ED 040 309. *Improving College Teaching Through Faculty Selection and Evaluation: A Review.* 4 pp. MF - 25¢; HC - 30¢.

ED 039 184. *Student Evaluation of Teaching and Learning.* 43 pp. MF - 25¢; HC - $2.25.

ED 036 485. *The Assessment Revolution: New Viewpoints for Teacher Evaluation.* 343 pp. MF - $1.50; HC - $17.25.

ED 035 416. *Identifying the Effective Instructor.* 34 pp. MF - 25¢; HC - $1.80.

ED 035 408. *Teacher Evaluation: Toward Improving Instruction.* 4 pp. MF - 25¢; HC - 30¢.

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE/Office of Education

National Center for Educational Communication